Creating AWS Job-Time Filesystems

A job-time filesystem is created at the beginning of each PBS batch job submitted from either a Pleiades front-end system (PFE) or an AWS dynamic front end. The filesystem is maintained while the job is running. When it is terminated at the end of the job, the data in the filesystem is lost. You are charged for the size provisioned for the lifetime of the job and the uptime of the filesystem server (if needed).

Note: Job-time filesystems are not created for interactive PBS jobs submitted from an AWS dynamic front end.

Adding the CLOUD Directives to Your PBS Script

Use the following directives in your PBS script to create various types of job-time filesystems, and to specify the size and mount point:

```
#CLOUD -volume_type=type
#CLOUD -volume_size=size_in_GB
#CLOUD -volume_mount=/path/to/mount/on
```

Important: The CLOUD directives must be placed *after* the the PBS directives in a PBS script.

-volume_type

The **-volume_type** directive must be placed before the **â volume_size** and **â volume_mount** directives in the PBS script; before **â volume_put** and **â volume_get**, which are described in this article; and before **â volume_primary**, which is described in this article.

See the next section for a list of options you can specify for the **â volume_type** directive.

-volume_size

For **-volume_size**, a number in gigabytes (G) is expected; any unit you add to it is ignored. For example, each of the following three specifications results in a request for a volume size of 10 G:

```
#CLOUD -volume_size=10
#CLOUD -volume_size=10G
#CLOUD -volume_size=10T
```

-volume_mount

For **-volume_mount**, mounting as **/home** is not currently allowed.

You can use as many sets of job-time filesystems as you want for each job. However, to reduce complexity for your job, we recommend using a maximum of two.

Types of Job-Time Filesystems

There are several types of job-time filesystems you can configure for your job. Some factors to consider when choosing a type are I/O pattern, size, performance, and cost.

You can specify the following options for the **â volume_type** directive:

- **ephemeral**

  Uses the ephemeral Non-Volatile Memory Express (NVMe)-based Solid State Drive (SSD) space that is local to various computing instance types. This type will ignore the **â volume_size** setting because the local space determines the size available.

  Not all instance types have a local disk. If the type you specify doesn't have disks available, it ignores the directive.

  For best results using NVMe-based SSD, add **arch=c5d** into the **select** line in the PBS script, as that is the only instance type currently supported by NAS. Note that **c5d** comes in six sizes, with 1, 2, 4, 8, 19, and 36 physical cores, and up to 1,800 GB of NVMeâ based SSD.
- **local**

  Each computing instance gets its own unique EBS filesystem. The size of the EBS volume unique to each computing instance is determined by **â volume_size=*x***. The *x* GB EBS volume accessed by one computing instance cannot be accessed by another computing instance, which has its own *x* GB of EBS to use.
- **headnode**

  Only the head node gets the EBS volumes. The size of the EBS is determined by the **â volume_size**. Other computing instances do not have access to the EBS on the head node.
- **node=*x***

  The *x*th computing instance gets an EBS filesystem. Specifying **node=0** has the same result as specifying **headnode**. The EBS filesystem of this *x*th instance is not accessible by another instance.
- **shared**

  Creates a distinct filesystem server instance that NFS-exports the volumes to all computing instances. None of the computing instances is used as the filesystem server.

  To select an instance as a filesystem server, it first looks at the **â volume_size** request. If there is an instance type that has at least the amount of local spaceâ NVMe SSD or Hard Disk Drive (HDD)â requested, that instance will be selected over other instances with EBS mount storage. Otherwise, a c5.18x instance with added EBS volumes to satisfy the required space will be selected.

  In other words, this "shared" filesystem has a few options for instances it will select behind the scenes. For the following requested volume sizes:

  - Under 1.8 TB: a c5d.18x instance with local SSD disks is selected.
  - Between 4 TB and 16 TB: an h1.16x instance with local HDD disk is selected.
  - For any other size: the normal c5.18x instance with EBS volumes is selected.
- **headnode**

  The EBS volumes on the head computing instance are NFS-exported to all other computing instances.

  Note: Exporting ephemeral space, if any, on the head computing instance to other computing instances is currently not supported by the NAS cloud developers.
- **lustre**

Creates a distinct filesystem using Amazon FSx for Lustre. The size will be determined by the value of **`-volume_size`**. There is an AWS-imposed minimum size of 3,600 GB. This filesystem will be mounted on all compute nodes at the mount point specified by **â `volume_mount`**.

Note: Using this filesystem will delay job startup by a few minutes while the filesystem is created.

For examples of how these directives are used, see this article.

Article ID: 591
Last updated: 25 Jan, 2021
Revision: 31
Cloud Computing -> AWS Cloud -> Creating AWS Job-Time Filesystems
https://www.nas.nasa.gov/hecc/support/kb/entry/591/